DOI: 10.1111/infa.12324

RESEARCH ARTICLE

INFANCY THE OFFICIAL JOURNAL OF THE INTERNATIONAL CONGRESS OF INFANT STUDIES WILEY

How 15-month-old infants process morphologically complex forms in an agglutinative language?

Enikő Ladányi[1,2,3] | Ágnes M. Kovács[4] | Judit Gervain[2,3]

[1]Department of Otolaryngology, Vanderbilt University School of Medicine, Nashville, Tennessee

[2]Integrative Neuroscience and Cognition Center (INCC–UMR 8002), CNRS, Paris, France

[3]Integrative Neuroscience and Cognition Center (INCC–UMR 8002), Université de Paris, Paris, France

[4]Department of Cognitive Science, Central European University, Budapest, Hungary

Correspondence
Enikő Ladányi, Department of Otolaryngology, Vanderbilt University School of Medicine, 1215 21st Avenue South, 37212 Nashville, TN, USA.
Email: eniko.ladanyi@vumc.org

Funding information
H2020 European Research Council, Grant/Award Number: 773202 ERC-2017-COG; Marie Sklodowska-Curie, Grant/Award Number: 641858

Abstract

While phonological development is well-studied in infants, we know less about morphological development. Previous studies suggest that infants around one year of age can process words analytically (i.e., they can decompose complex forms to a word stem and its affixes) in morphologically simpler languages such as English and French. The current study explored whether 15-month-old infants learning Hungarian, a morphologically complex, agglutinative language with vowel harmony, are able to decompose words into a word stem and a suffix. Potential differences between analytical processing of complex forms with back versus front vowels were also studied. The results of Experiment 1 indicate that Hungarian infants process morphologically complex words analytically when they contain a frequent suffix. Analytic processing is present both in the case of complex forms with back and front vowels according to the results of Experiment 2. In light of the results, we argue for the potential relevance of the early development of analytic processing for language development.

1 | INTRODUCTION

While English uses three words to say *in my house*, the same is expressed with a single word *házamban* in an agglutinative language like Hungarian. The stem *ház* "house" takes the 1st person possessive suffix *–am* and the locative case marking *–ban*. In agglutinative languages, nouns and verbs typically appear in such morphologically complex forms. How can infants meet the challenge of learning to decompose these forms and extracting the word stems? The current paper seeks to answer this question by investigating whether 15-month-old Hungarian infants can extract a nonce noun stem when it is followed by a frequent suffix (*–ban*/*–ben* "in").

Affixes, also known as bound morphemes, are grammatical morphemes appearing before or after (prefixes and suffixes, respectively) the word stem, constituting a morphologically complex word form with the stem. Languages vary systematically in the complexity of the affixation they allow. Traditional descriptive linguistics differentiates between four categories (Sapir, 1921; Humboldt, 1836/1999; see also Comrie, 1980; Baker, 2010). Isolating languages such as Mandarin Chinese have very little affixation (e.g., *wɔ mən tan tçin lə*: 1st person.pl.play.piano.past "we played the piano"). In fusional languages, grammatical functions are expressed with a small set of bound morphemes in addition to free morphemes, and one morpheme marks several functions (e.g., Spanish *hablo* "I am speaking," the suffix *-o* means 1st person, singular and present tense). In agglutinative languages, grammatical relations are expressed with a large set of affixes, and one grammatical function is typically marked by one morpheme (e.g., Hungarian *zongorá-k-hoz* piano.pl.to "to the pianos"). Polysynthetic languages have both agglutinative and fusional morphemes, and they can also have multiple stems in a single word by incorporating the subject and object nouns into the verb (e.g., Sora *anin- ɲam-jɔ-te-n* he.catch.fish.pres.aux "he is fish-catching").

The combination of word stems and affixes is governed by (semi-)productive rules. As a result, we are able to produce and comprehend complex forms never heard before if we are familiar with the constituents. Experimental results suggest that the word stem and the affix(es) of complex forms are automatically activated in adults when they are exposed to a morphologically complex form (Caramazza, Laudanna, & Romani, 1988; Schreuder & Baayen, 1995). Beyond enabling the processing of novel word forms, analytic mechanisms are more economic than storing each morphologically complex form as a separate entry. At the same time, however, analytic processing is computationally more demanding than holistic processing.

Can infants also represent complex forms analytically? The development of analytic processing is especially crucial in languages with a rich morphology, as the successful decomposition of complex forms into its constituents is necessary for word learning. We are not aware of, however, any studies exploring the representation of morphologically complex forms in infants acquiring a morphologically complex language. In this study, we therefore aim to fill a gap in the literature by examining whether infants learning a morphologically rich agglutinative language can decompose complex forms. (We are only focusing on inflectional morphology in this study, leaving aside derivational morphology, that is, affixation that changes the meaning and/or the lexical category of a stem, turning it into another stem, for example, *beauty* + *-ful* -> *beautiful*.)

1.1 | Hungarian morphology

In the current study, we target Hungarian, an agglutinative language, a member of the Uralic/Finno-Ugric language family. Since grammatical relations are expressed by suffixes, word order is relatively free in Hungarian. There are a large number of nominal and verbal suffixes, and a stem can be followed by several suffixes. Nouns, for instance, can be followed by plural and possessive markers as well as a case marker (e.g., *házainkban* ("in our houses") in which *ház* ("house") is the word stem, *-aink* ("our") is a possessive marker, and *-ban* ("in") is a case marker). The order of these suffixes is fixed. As a result of the large number of possible suffixes, for example, 18 different case markers, a noun can have more than 700 different forms (Gábor & Lukács, 2012). Verb morphology is also similarly complex. For instance, *olvashatnád* ("if you may read something") consists of the word stem *olvas* ("read"), the modal suffix *-hat* and *-nád*, the 2nd person singular definite verb suffix in present tense and conditional mood.

Similarly to many other agglutinative languages, Hungarian shows vowel harmony. Vowels within a word are homogeneous with respect to one or several phonological features. In contrast to some other harmonic languages (e.g., Turkish), vowel harmony is not fully regular in Hungarian: According to corpus analyses of child-directed speech, approximately 70% of Hungarian words show vowel harmony, while for instance in Turkish 85%–90% do (Gonzalez-Gomez, Schmandt, Fazekas, Nazzi, & Gervain, 2019; Ketrez, 2014). The most prevalent phonological feature that requires harmony in Hungarian is the backness of a vowel. Most of the words have either only front vowels (e.g., *szekrény* "cupboard") or only back vowels (e.g., *barlang* "cave"). Importantly for our research question, harmony also appears at the morphological level with word stems and their affixes harmonizing in backness. In order to satisfy vowel harmony, most suffixes have two allomorphs: one with a front vowel and one with a back vowel. For instance, the inessive suffix *-bAn* ("in") is realized as *-ben* following the word *szekrény* (*szekrényben* "in the cupboard"), but as *-ban* following the word *barlang* (*barlangban* "in the cave"). (The standard notation to refer to the two allomorphs together is *-bAn*; see more details about vowel harmony in Hungarian in Gonzalez-Gomez et al., 2019 and Törkenczy, 2011.)

1.2 | The development of morphology

Research targeting morphological acquisition in Hungarian, as in most other agglutinative languages, is limited. There are a few early studies discussing the use of bound morphemes in the spontaneous speech of children acquiring Hungarian (e.g., Mikes, 1967; Slobin, 1973; Slobin, 1982). These studies are crucial because while research focused on English for a long time, they drew attention to the different developmental trajectories in an agglutinative language. However, at the same time, by studying spontaneous speech samples, it is impossible to explore the question of analytic processing since the production of suffixed forms can be explained by both holistic and analytic processes/encoding. A few studies targeted production and have explored whether children are able to use bound morphemes in a productive way. In these studies, children learned a pseudo-word in the first part of the experiment, which they had to use in a suffixed form in the second part (somewhat similarly to the classical *wug*-test, Berko, 1958). The results of these studies indicate that children can produce the suffixed form of a pseudo-word at age three (MacWhinney, 1975, 1978) or even earlier, at two and a half years for nouns (Gábor & Lukács, 2012). These findings suggest that children have a representation of suffixes and are able to apply morphological rules in a productive way to create morphologically complex forms at this age. Research on morphological comprehension and analysis is even more limited. To our knowledge, no studies have explored when Hungarian infants start to represent bound morphemes as distinct units, and recognize the word stem and the suffix as separate morphemes in the speech input.

Research concerning the development of analytic processing in other morphologically less complex languages is also scarce. We are aware of only two studies (Marquis & Shi, 2012; Mintz, 2013) directly targeting this question. The results of these studies suggest that infants acquiring French or English are able to process suffixed forms analytically at around one year of age.

Marquis and Shi (2012) investigated 11-month-old Canadian French-learning infants. In their study, infants were repeatedly presented with an isolated pseudo-word (/trid/) in the first part of the experiment. In the second part, either the same pseudo-word or another pseudo-word (/glyt/) was presented together with a frequent verbal suffix (/-e/) in sentences like *Maman a/**tride**/les poèmes*

(“Mum **trided** the poems”). The authors measured listening times to the two types of sentences (with the target vs. the control pseudo-word). If infants already represent /-e/ as a suffix and process complex forms analytically, they will recognize the familiarized word stem. Listening times were found to be longer in the target condition, suggesting that infants were able to process the suffixed form analytically. However, infants might have simply showed longer listening times for the target condition because there is a partial word-form overlap between the target form and the suffixed form *(/trid/–/tride/*). To exclude this possibility, another group of infants was presented with the same word stems followed by a pseudo-suffix *-ou* (/-u/) during the test phase (*Maman a* ***/tridu/*** *les poèmes*). In this condition, listening times did not differ between target and control conditions, suggesting that phonetic similarity is not enough for a familiarity effect, and that in the original experiment, the familiarity effect appeared due to segmentation and not to partial word-form overlap. The authors concluded that infants were able to decompose the morphologically complex form in the first condition because they had already extracted */–e/* as a suffix from the speech they were exposed to in their everyday life before they came to the laboratory. They could do it because in the natural input */–e/* frequently appears with various word stems. In contrast, */–u/* is not a verbal suffix in French, which is why they did not extract it in the second condition.

Interestingly, in a further experiment Marquis and Shi (2012) also tested whether infants can learn */–u/* as a new suffix and whether they can use this newly learned pseudo-suffix to process the morphologically complex form in the test phase analytically. In Experiment 2, the authors presented infants with the pseudo-suffix */-u/* together with different pseudo-stems (e.g., *linchou*, *cradou*, *plandou*, *wélou…*) before the familiarization phase. This learning phase gave infants the possibility to extract */-u/* as a suffix. After the learning phase, they were presented with the same familiarization and test phase as in the pseudo-suffix condition of Experiment 1. This time infants showed the same effect as in the real suffix condition of Experiment 1. These results suggest that French-learning 11-month-old infants are able to extract suffixes from the speech stream based on their distributional properties (i.e., a small set of suffixes appear frequently together with different word stems) and can use these previously encoded suffixes to analyze the internal structure of newly encountered words.

We note that the syntactic structures of the sentences used in this study support an analysis of the complex pseudo-words as suffixed forms (e.g., *Maman a /tridu/…*, where */tridu/* is preceded by an auxiliary that takes a participle form of the verb to form the past tense). The sentences would be ungrammatical if the pseudo-word was analyzed holistically, as a non-suffixed noun/verb. This might thus affect the segmentation process.

Mintz (2013) raised some methodological considerations with respect to the above results. As he argues, a more robust measure of segmentation abilities would be a design in which infants are presented with the suffixed form of the word stem in passages during the familiarization phase, and then listen to bare stems during test. This design is more demanding both for memory and for segmentation, but importantly, it is more similar to the task infants have to accomplish during real-life language acquisition.

In accordance with these arguments, Mintz (2013) conducted segmentation experiments with 15-month-old English-learning infants. In two experiments, infants were presented first with the suffixed form of pseudo-words in sentences. During the familiarization phase, infants were presented with two pseudo-words with a real suffix (*-ing*: *I see you* ***lérjoving****! Johny likes* ***gemónting****!*) and with two pseudo-words with a pseudo-suffix (*-dut*: *Does Sam want to go* ***káfteedut****? I want to go* ***jivántdut****!*). In the test phase, infants listened to one of the bare word stems repeatedly in each condition. The result that infants showed longer listening times for the word stems, which were presented together with the real suffix (*-ing*) than to those, which appeared with the pseudo-suffix (*-dut*), suggests that 15-month-old English-speaking infants process morphologically complex forms analytically. A further experiment did not show any evidence for analytic processing in 8-month-olds in the same paradigm.

This study eliminates the problem of the role of the syntactic context by using sentences that are syntactically correct with both a suffixed and a non-suffixed word form (e.g., *Does Sam want to go* ***káfteedut****? – kaftedut* can be replaced by both a suffixed form like *fishing* and a non-suffixed form like *home).*

In sum, the above studies suggest that 11-month-old French- and 15-month-old English-learning infants are able to process morphologically complex forms analytically.

In the current study, we aimed to investigate whether 15-month-old Hungarian-learning infants are able to process morphologically complex forms analytically. On the one hand, complex morphology might facilitate analytic processing, since the large number of morphologically complex words might be too demanding to store holistically. On the other hand, the high variability of these forms and the existence of allomorphs might make the task more challenging and delay the start of analytic processing. In Experiment 1, we, therefore, explore whether 15-month-old Hungarian infants process pseudo-words with a suffix analytically. In Experiment 2, potential differences between the acquisition of vowel harmonic allomorphs (front vs. back vowel) of the suffix will be examined.

2 | EXPERIMENT 1

In Experiment 1, we aimed to test whether 15-month-old Hungarian infants represent the inessive suffix *-bAn* ("in") and are able to process forms with this suffix analytically. The inessive suffix has two allomorphs: *-ban*, including a back vowel, and *-ben*, including a front vowel. Since the decomposition of morphologically complex forms might depend on the frequency of exposure to the specific allomorph of the suffix (see Shi, Cutler, Werker, & Cruickshank, 2006; Shi, Lepage, Gauthier, & Marquis, 2006), we investigated the frequency of the front and the back versions of the suffix *-bAn* in two corpora. To explore frequencies in child-directed speech, we calculated the token frequencies of nouns with *-ban* and with *-ben* in the adult utterances of the Hungarian subcorpora of the CHILDES corpus (MacWhinney, 2000). Given to the limited size of this corpus, we also queried the token frequencies of words with the two allomorphs in a WebCorpus (*Szószablya* database, Németh, 2003). The two corpora showed the same pattern: The frequency of the front vowel allomorph was somewhat lower than the frequency of the back vowel allomorph although the difference was smaller in the larger adult corpus (see frequencies in Table 1).

Motivated by the higher frequency of the back vowel allomorph, we decided to use this allomorph in Experiment 1 to test whether infants are able to decompose morphologically complex forms. In Experiment 2, we then investigated potential differences between the two allomorphs.

In Experiment 1, a group of infants was presented with a pseudo-word (e.g., *púr*) with the suffix *-ban* in simple Hungarian sentences during the familiarization phase. During the test phase, listening times for the familiarized pseudo-root and a control pseudo-root (e.g., *gál*) were measured. Similarly to Mintz (2013), we presented infants with sentences during the familiarization phase and measured listening times to the bare pseudo-word stems during the test phase. While more challenging, this

TABLE 1 Token frequencies of words with suffix *-ban* and *-ben* in the CHILDES and Szószablya corpora

	-ban	-ben
CHILDES	357	250
Szószablya	~3.8 million	~3.4 million

TABLE 2 Stimuli of Experiment 1 in the four groups of infants

Group	Suffix	Target word	Familiarization: Example sentences	Test phase
1 ($n = 14$)	Real suffix	púr	*Púrban lakni jó!* "It's great to live in a púr!" *Az egyik púrban sötét van.* "It's dark in one of the púrs" *Elolvastam a könyvet* a *púrban.* "I read the book in the púr"	target: *púr, púr, púr…* control: *gál, gál, gál…*
2 ($n = 12$)		gál	*Gálban lakni jó!* *Az egyik gálban sötét van.* *Elolvastam a könyvet* a gálban.	target: *gál, gál, gál…* control: *púr, púr, púr…*
3 ($n = 13$)	Pseudo-suffix	púr	*A púrdag nézegeti a virágokat* "The púrdag is watching the flowers"/"He is watching the flowers in the púr" *Kényelmesebb a hűvös púrdag.* "The chilly púrdag is more comfortable"/"It is more comfortable in the chilly púr"	target: *púr, púr, púr…* control: *gál, gál, gál…*
4 ($n = 13$)		gál	*A gáldag nézegeti a virágokat.* *Kényelmesebb a hűvös gáldag.*	target: *gál, gál, gál…* control: *púr, púr, púr…*

method is closer to real-life segmentation, as also argued by Mintz (2013). As we used Mintz's (2013) paradigm, we decided to test infants of the same age as in his study.

Since unlike *púrban* and *gál*, *púrban* and *púr* sound similar, infants might show listening time differences between *púr* and *gál* without the decomposition of the complex form *púrban*, simply as a result of similarity in the word forms. To control for the effect of partial word-form overlap, another group of infants listened to the same pseudo-word with a pseudo-suffix *-dag* and presented them with the same testing phase as the first group of infants.

We expected 15-month-old Hungarian infants to differentiate between the familiarized and the non-familiarized forms when they were presented together with the real suffix *–ban,* but not when they appeared with a pseudo-suffix *-dag* during the familiarization.

2.1 | Methods

2.1.1 | Participants

Fifty-two 15-month-old monolingual Hungarian-learning infants participated in the study (23 girls, mean age: 15 months 26 days, *SD*: 12 days, range: 15 months 6 days–16 months 16 days). Of these, 14 infants were familiarized with the pseudo-word *púr* and the real suffix *–ban*, 12 infants with the pseudo-word *gál* and the real suffix *–ban*, 13 with the pseudo-word *púr* and the pseudo-suffix *–dag*, and 13 with the pseudo-word *gál* and the pseudo-suffix *–dag*. An additional 31 infants were tested but excluded from the analysis due to completing less than one trial per block per target/control condition ($n = 14$), fuzziness/crying ($n = 16$) or parental interference ($n = 1$).

Infants participating in all experiments reported in this paper were recruited from the infant database of the Central European University. The present study was conducted according to guidelines laid down in the Declaration of Helsinki, with written informed consent obtained from a parent or guardian for each child before any assessment or data collection. All procedures involving human subjects in this study were approved by the United Ethical Review Committee for Research in Psychology, Hungary, at the Central European University.

2.1.2 | Stimuli and design

Two pseudo-words following the phonotactic rules of Hungarian were created with a CVC structure (*púr*, *gál*). For the familiarization phase, these pseudo-words were inserted into Hungarian sentences with a simple syntactic structure and composed of high frequency words (Table 2). For two groups of infants, each of the two pseudo-words appeared with the existing Hungarian suffix *–ban*; for another two groups of infants, they appeared with the pseudo-suffix *-dag*. We chose the suffix *–ban*, because it is highly frequent, it is a full syllable, and it is thus phonologically sufficiently salient (as opposed to single consonant suffixes, such as the accusative case marker *–t*, which are harder to perceive). The pseudo-suffix *-dag* has a phonological structure similar to the real suffix *–ban*, but it is not a suffix in Hungarian. At the same time, it is a legal word ending in Hungarian.

The suffixed form of the pseudo-word appeared at the beginning, the middle, or the end of the sentences. Importantly, we created sentence contexts for the pseudo-suffixed form that are grammatical both when the form is parsed as a word stem (*púrdag: A púrdag nézegeti a virágokat.* "The púrdag is looking at the flowers.") and as a suffixed noun (*púr + dag: A púrdag nézegeti a virágokat.* "He/she is looking at the flowers in/on/up… the púr."). The pseudo-words appeared in 36 sentences; therefore, altogether four sets of 36 sentences were created for the familiarization phase (one set for the familiarization phase for each group of infants).

A female, native Hungarian speaker naïve to the aim of the study recorded the stimuli in a sound-attenuated room in a child-directed fashion (sampling rate 44.1 KHz, bit rate 16 bits). First, she recorded the 36 sentences four times with the four different pseudo-words. Then, several instances of each pseudo-word were recorded without the suffix for the test phase. For the familiarization phase, the 36 sentences belonging to one condition were concatenated into a continuous stream with a 1-s pause between each sentence. The total lengths of the familiarization streams were 2 m 4 s for *púrban*, 2 m 8 s for *gálban*, 2 m 2 s for *púrdag*, and 2 m 4 s for *gáldag*. For the test phase, we selected 15 instances of both pseudo-words on the basis of recording quality and we concatenated them with 500 ms pauses. The total length of the test stimuli was 15.33 s for *púr* and 16.99 s for *gál*.

2.1.3 | Procedure

The head-turn preference procedure (HPP; Saffran, Johnson, Aslin, & Newport, 1999) was used to test whether infants can differentiate between a pseudo-word that they heard together with a suffix during the familiarization phase and another pseudo-word that they never heard before. Since the Cognitive Development Center of the Central European University, where infants were tested, is not equipped with a classical three-light HPP setup, a large, 40-inch plasma TV screen was used to measure listening times at three different locations. This setup was successfully used in another speech perception-related HPP task in this laboratory (Gonzalez-Gomez et al., 2019). The screen was located in a dimly lit, soundproof room. Auditory stimuli were presented from two loudspeakers hidden in the left and right side of the room. During the experiment, the infant was sitting in her parent's lap approximately 36 inches away from the screen. Parents were listening to masking music through headphones during the experiment to avoid involuntary influence on the infant's behavior. They were also asked to avoid any interaction with the infant during the experiment. A camera was used to project the infant's looking behavior to a screen outside of the room. Stimuli were presented with the experimental software PsyScope X (Cohen, MacWhinney, Flatt, & Provost, 1993). Presentation of stimuli was controlled by an experimenter, seated outside the booth, and blind to the auditory stimuli, solely based on the looking behavior of the child. The PsyScope software recorded the experimenter's button presses indicating infant's looking behavior online and calculated look away. Recordings of the infants' looking behavior were kept for subsequent offline coding, and these data were used in the data analysis instead of the looking behavior measured online, as offline coding provides higher precision.

First, infants listened to the approximately 2-min-long familiarization stream with the sentences presented from both loudspeakers. The sound was not contingent on infants' looking behavior. While infants were listening to the auditory stimuli, a video animation depicting a green flashing light appeared at the center of the screen (central light). If the infant fixated on it, a yellow light started to blink on one side of the screen (side light). The side light was on as long as infants fixated on it. When infants turned away from the light and did not turn back within two seconds (measured by the software), the side light was extinguished, and the central light appeared again.

After the familiarization, infants were presented with two blocks of four test trials. At the beginning of each trial, the central light started to blink and after the infant fixated on it, it was extinguished and one of the side lights started to blink. When the infant fixated on the side light, the auditory stimulus started to be played from the loudspeaker on the side of the light. It continued until the end of the trial, that is, until the 15 repetitions of the same test item were presented, or until the infant looked away for more than two seconds. Infants were presented four times with the word of the target condition and four times with the word of the control condition. The presentation order of the stimuli was pseudo-randomized with the constraint that the same condition was presented consecutively at most twice and that both conditions were presented twice in the first block and twice in the second block. Presentation side of the stimuli (i.e., the side of the screen on which the light appeared together with the side of the loudspeaker from which the auditory stimulus was presented) was pseudo-randomized and counterbalanced across infants. A stimulus could not appear more than twice consecutively on the same side.

Infants were randomly assigned to one of the four conditions. Two groups of infants heard the pseudo-word with the existing suffix *-ban* during the familiarization phase, and one of these groups was familiarized with *púrban*, while the other group was familiarized with *gálban*. The other two groups were presented with the pseudo-word with the pseudo-suffix *-dag* (one group with *púrdag* as the target item and the other with *gáldag* as the target item). The design was thus counterbalanced across infants: The target trials for one group served as the control trials for the other group and vice versa.

Listening times were offline coded by the experimenter, and half of the videos were also coded by a second coder. The intercoder reliability was 98.5%.

2.1.4 | Data analysis

The criteria for trial inclusion were to have listening times of at least 1 s to make sure infants heard at least one instance of the stem. Mean listening times were analyzed with a $2 \times 2 \times 2 \times 2$ ANOVA with suffix (real suffix vs. pseudo-suffix) and familiarized word (púr vs. gál) as between-subject factors and familiarity (target vs. control) and block (first vs. second) as within-subject factors.

2.2 | Results

The average listening times in the real suffix and pseudo-suffix conditions are shown in Figure 1. The 2 × 2 × 2 × 2 ANOVA showed a main effect of block ($F(1, 48) = 14.789$, $p < .001$) as a result of longer listening times in the first block (mean = 11,093 ms, $SD = 3{,}515$ ms) than in the second block (mean = 9,401 ms, $SD = 3{,}864$ ms) of the test phase. The main effects of the suffix ($F(1, 48) = 4.312$, $p = .043$) and the familiarized word ($F(1, 48) = 5.276$, $p = .026$) were also significant. Infants who were familiarized with *púr* showed generally longer listening times during the test phase (mean = 10,941 ms, $SD = 3{,}662$ ms) than infants who were familiarized with *gál* (mean = 9,497 ms, $SD = 3{,}781$ ms). Infants who were familiarized with the pseudo-suffix *-dag* (mean = 10,893 ms, $SD = 3{,}718$ ms) also showed longer listening times than infants who were familiarized with the real suffix *-ban* (mean = 9,601 ms, $SD = 3{,}749$ ms). The interaction between familiarity and block ($F(1, 48) = 6.077$, $p = .017$) and, importantly, between familiarity, block, and suffix ($F(48, 1) = 5.340$, $p = .025$) were also significant. To explore these interactions further, we conducted a 2 × 2 × 2 ANOVA separately for the real suffix and the pseudo-suffix conditions. In the pseudo-suffix condition only, the main effect of block was significant ($F(1, 24) = 8.306$, $p = .008$). In the real suffix condition, the block × familiarized word interaction ($F(1, 24) = 9.850$, $p = .004$) was also significant beyond the main effect of block ($F(1, 24) = 6.526$, $p = .017$). According to follow-up paired sampled t tests, the interaction appeared because infants listened longer to the control item than to the target in the first block ($t(25) = 2.661$, $p = .013$) but no significant difference appeared in the second block of the testing phase ($t(25) = 1.952$, $p = .062$).

2.3 | Discussion

The statistical analyses yielded several significant main effects showing the usual decrease in listening times throughout the experiment, longer listening times in the groups of infants who were exposed to *púrban* or *púrdag* than in those who were listening to *gálban* or *gáldag* during the familiarization phase and listening times were also longer in infants familiarized with *púrdag* or *gáldag* than in those who were familiarized with *púrban* or *gálban*. Beyond these general tendencies, which we come back to below, importantly, we found that in the first block infants looked longer to the control pseudo-word over the target pseudo-word when the real suffix was used, but not when the pseudo-suffix was

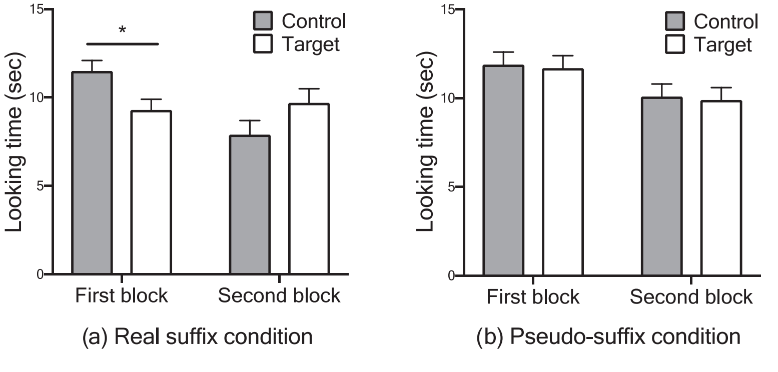

FIGURE 1 Mean listening times (with standard errors) in Experiment 1 to the control and target conditions in the first and the second block in (a) the real suffix condition and (b) the pseudo-suffix condition (collapsed across familiarized word). Significant differences at $p < .05$ are marked with *

used. In the second block, this difference was no longer present, probably due to the overall decreasing listening times and/or habituation to the control stem.

Difference in listening times between the target and control pseudo-word in the real suffix, but not in the pseudo-suffix condition suggests that 15-month-old Hungarian infants successfully use their prior knowledge of the real suffix to decompose morphologically complex word forms and extract the stem. This is in accordance with earlier results in 11-month-old French-learning (Marquis & Shi, 2012) and 15-month-old English-learning infants (Mintz, 2013, Experiments 2 and 3). At the same time, these two studies found a familiarity effect, while in our study, a novelty effect was observed. According to the attention model of Hunter and Ames (1988) if in a looking time paradigm the familiarization phase is short or the task is difficult, infants will show a familiarity effect as a sign of ongoing, incomplete learning/processing, whereas if the familiarization is longer or the task is less difficult, then a novelty effect will appear in the testing phase, as infants will have completed processing the stimuli by the end of the familiarization phase. We speculate that Hungarian-learning infants might have shown a novelty effect due to the richer morphology of their mother tongue making analytic processing less demanding for them than for French- or English-learning infants.

Another difference between the results of the current study and previous studies is that we found a listening time difference only in the first block, while previous studies found it in the whole test phase. One potential explanation is that, given their greater experience with a richer morphology, the task might have been less difficult for the Hungarian infants, and did not pay much attention in the second block. The main effect of block due to shorter listening times in the second block than in the first block does indeed suggest a drop in attention during the testing phase.

Two other main effects were also significant. Listening times were longer in infants familiarized with *púr* (*púrban* or *púrdag*) than in infants familiarized with *gál* (*gálban* or *gáldag*) and in infants who were familiarized with the pseudo-suffix than in those who were exposed to the real suffix. These effects are probably related to the idiosyncratic properties of the stimuli and did not interact with familiarity, the factor of interest in our study.

In sum, results of Experiment 1 suggest that Hungarian-learning 15-month-old infants are able to process morphologically complex forms analytically. Hungarian, however, has vowel harmony. The suffix investigated in Experiment 1 thus also has an allomorph with a front vowel: *-ben* which is slightly less frequent than the back vowel allomorph (*-ban*), used in Experiment 1. In Experiment 2, we explored whether complex forms with the front vowel allomorph are also processed analytically by 15-month-old infants to replicate the results of Experiment 1, and to test the generality of decomposition. With Experiment 2, we can also investigate whether the slight frequency difference between the two allomorphs has an effect on analytical processing.

3 | EXPERIMENT 2

3.1 | Methods

3.1.1 | Participants

Twenty-seven 15-month-old monolingual Hungarian-learning infants participated in the study (14 girls, mean age: 15 months 24 days, *SD*: 12 days, range: 15 months 2 days–16 months 15 days). Of these, 14 infants were familiarized with sentences containing *pérben* and 13 with sentences containing *gőlben*. An additional 11 infants were tested but excluded from the analysis due to completing less than one trial per block per condition ($n = 9$) or fuzziness/crying ($n = 2$).

3.1.2 | Stimuli and design

We created two pseudo-words, which shared their consonants with the pseudo-words of Experiment 1, but had front rather than back vowels: *pér* and *gől*. Words were inserted in the same 36 sentences as in the real suffix condition of Experiment 1 together with the front vowel allomorph of the suffix, *-ben* (Table 3). The recording and the production of the stimuli were identical to Experiment 1. The total lengths of the familiarization streams were 2 m 9 s for the stream with *pérben* and 2 m 4 s for the stream with *gőlben*. The total lengths of the test trials were 15.10 s for *pér* and 16.18 s for *gől*.

3.1.3 | Procedure

The procedure was identical to Experiment 1. All the videos were coded by a second coder with an intercoder reliability of 98.4%.

3.1.4 | Data analysis

To investigate whether analytic processing of words with the suffix *-bAn* is also present in the case of the front vowel allomorph of the suffix, we analyzed the data of this experiment together with the corresponding data of the real suffix condition of Experiment 1. We thus ran a 2 × 2 × 2 × 2 ANOVA with vowel type (back vowel vs. front vowel) and consonant tier (p_r tier: púr, pér vs. g_l tier: gál, gől) as between-subject factors and familiarity (target vs. control) and block (first vs. second) as within-subject factors.

3.2 | Results

Mean listening times analyzed in Experiment 2 are shown in Figure 2. The 2 × 2 × 2 × 2 ANOVA showed the main effect of block ($F(1, 49) = 35.088$, $p < .001$), due to significantly longer listening times in the first block than in the second block of the test trials. The interaction between block and vowel type ($F(1, 49) = 5.052$, $p = .029$) and between block and familiarity ($F(1, 49) = 9.469$, $p = .003$) was also significant. According to follow-up ANOVAs, in the first block, infants showed marginally longer listening times in the front vowel condition than in the back vowel condition ($F(1, 52) = 3.094$, $p = .085$), while in the second block, there was no such difference ($F(1, 52) = 1.097$, n.s.). Further

TABLE 3 Stimuli of Experiment 2 in the two groups of infants

	Familiarization	Test phase
pér	*Pérben lakni jó!* “It's great to live in a pér!” *Az egyik pérben sötét van.* “It's dark in one of the pérs” *Elolvastam a könyvet a pérben.* “I read the book in the pér”	target: *pér, pér, pér…* control: *gől, gől, gől…*
gől	*Gőlben lakni jó!* *Az egyik gőlben sötét van.* *Elolvastam a könyvet a gőlben.*	target: *gől, gől, gől…* control: *pér, pér, pér…*

analysis of the block x familiarity interaction showed significantly longer listening times for the control pseudo-word than for the target pseudo-word in the first block ($t(53) = 3.358$, $p = .001$), but not in the second block of the test phase ($t(53) = 1.033$, n.s.).

3.3 | Discussion

In Experiment 2, we found that infants showed longer listening times to the control words in the first block and that backness did not have an effect on this effect. These results suggest that infants are able to process morphologically complex forms both with front and back vowels analytically at 15 months of age despite the slight difference in the frequencies of the two harmonic allomorphs.

4 | GENERAL DISCUSSION

Our study investigated whether 15-month-old Hungarian-learning infants are able to process morphologically complex forms containing the frequent locative suffix *-bAn* ("in") analytically. In Experiment 1, we found that infants looked longer to the control word to the target word in the first block of the testing phase when the target word was presented with the suffix *-ban* during familiarization. At the same time, no listening time difference appeared when the target word was familiarized with the pseudo-suffix *-dag*. These results indicate that infants were able to process the pseudo-word with a real suffix analytically during the familiarization phase and could thus differentiate between target and control trials during the testing phase. The null effect in the pseudo-suffix condition suggests that the novelty effect did not appear due to the partial word-form overlap between the suffixed form and the word stem. In Experiment 2, we showed that the backness of the vowels in the word stem and the suffix did not interact with the novelty effect suggesting that infants were able to decompose words with both the front vowel and back vowel allomorphs despite the somewhat higher frequency of the back allomorph.

To our knowledge, these are the first results that show morphological decomposition in an agglutinative language early in development. As discussed above, this finding is in line with earlier studies' finding evidence for decomposition in French- and English-learning infants around one year of age (Marquis & Shi, 2012; Mintz, 2013). We interpreted the differences between our study and these

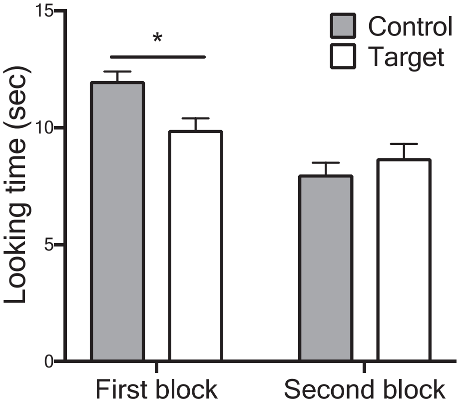

FIGURE 2 Mean listening times (with standard errors; collapsed across Experiment 1 and Experiment 2) to the control and target conditions in the first and the second block (collapsed across vowel type and consonant tier). Significant differences at $p < .05$ are marked with *

previous works—novelty instead of familiarity effect and difference only in the first block and not in the whole testing phase—as indicators of potentially more advanced morphological decomposition in Hungarian-learning infants than in English- or French-learning infants due to the richer morphology of Hungarian. More advanced morphological decomposition in Hungarian-learning infants might originate from the fact that Hungarian word stems can have hundreds of different suffixed forms and their holistic storage might become too demanding for infants' memory, and thus trigger analytic processing. Morphological complexity might thus not be a hindrance, but instead a cue that may trigger the appropriate strategies, such as decomposition, that allow efficient learning. However, comparisons between previous studies and the current study should be handled with caution due to the differences in their experimental designs.

It is important to note that although our study suggests that infants are able to process words with the suffix *-bAn* analytically, further studies will need to test whether these results are generalizable to other suffixes. The suffix *-bAn* has a salient CVC structure, and it is one of the most frequent suffixes; therefore, it is probably among the first ones that infants can segment from continuous speech.

Our results suggest that 15-month-old infants have a representation of the sound form of the suffix *-ban* and *-ben*, but they cannot provide any information about whether they know the meaning of the suffix. As discussed earlier, Marquis and Shi (2012) found evidence for 11-month-old French infants' ability to learn a novel pseudo-suffix form without any meaning, indicating that infants might first segment the sound form of suffixes, and only later attach meaning to them. This may also be modulated by the semantic contents of the suffix.

Since many Hungarian suffixes including a vowel have at least two allomorphs to satisfy vowel harmony, we also explored potential differences between the analytic processing of morphologically complex forms with a front and a back vowel. According to our corpus analyses, the back vowel allomorph of the suffix *-bAn* is slightly more frequent than the front vowel allomorph. Our study did not find any difference between the two allomorphs; however, it does not exclude the possibility that there is a developmental difference between the two allomorphs earlier in language acquisition.

Our results can be connected to the line of literature targeting the acquisition of free function words. Research about infants' knowledge of function words is more extended than research about bound morphemes. These studies found that infants are sensitive to perceptual (acoustic and phonological) differences between function words and content words already from birth (Shi, Werker, & Morgan, 1999). Infants have also been shown to have a representation of free functional morphemes from as early as 7–11 months (e.g., Hallé, Durand, & de Boysson-Bardies, 2008; Höhle & Weissenborn, 2003; Shafer, Shucard, Shucard, & Gerken, 1998; Shi, Cutler, et al., 2006; Shi, Marquis, & Gauthier, 2006; Shi, Werker, & Cutler, 2006), and they use these elements in several processes, like word segmentation (Shi, Cutler, et al., 2006; Shi & Lepage, 2008) or grammatical categorization (Gerken, Wilson, & Lewis, 2005; Höhle, Weissenborn, Kiefer, Schulz, & Schmitz, 2004; Mintz, 2006; Shi & Melançon, 2010). We speculate that bound morphemes have similar functions for young infants—a hypothesis that will require empirical conformation in the future.

We believe that our study is an important first step in exploring the development of analytic processing in agglutinative languages. Further studies with typologically different languages will be needed to determine whether the development of analytic processing depends on the ambient language.

5 | CONCLUSION

Our study is the first to show evidence of analytic processing in infants in an agglutinative language. Fifteen-month-old Hungarian infants were found to be able to process morphologically complex forms with both front and back vowels analytically. Future studies should explore whether younger infants are also able to process morphologically complex forms analytically and investigate the relationship between analytic processing and grammatical development.

ACKNOWLEDGMENTS

This study was conducted with the support of the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie Grant Agreement No. 641858 (PredictAble project) to J.G and E.L and the ERC Consolidator Grant (773202 ERC-2017-COG "BabyRhythm"; https://erc.europa.eu/funding/consolidator-grants) to J.G. The authors declare no conflicts of interest with regard to the funding source for this study. Special thanks go to the infants and their parents for their kindness and cooperation. We thank Peter Rebrus for his help in creating experimental stimuli and Agnieszka Argasińska for the offline coding of the video recordings. We thank Gergely Csibra, Ágnes Volein, Mari Tóth and the whole Cognitive Developmental Center of the Central European University, Budapest, Hungary, for their invaluable help and support with testing infants.

ORCID

Enikő Ladányi https://orcid.org/0000-0003-2853-682X

REFERENCES

Baker, M. (2010). Formal generative typology. In B. Heine, & H. Narrog (Eds.), *The Oxford handbook of linguistic analysis* (pp. 285–312). Oxford, UK: Oxford University Press.

Berko, J. (1958). The child's learning of English morphology. *Word*, *14*(2–3), 150–177. https://doi.org/10.1080/00437956.1958.11659661

Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition*, *28*(3), 297–332. https://doi.org/10.1016/0010-0277(88)90017-0

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers*, *25*(2), 257–271.

Comrie, B. (1980). *Language universals and linguistic typology: Syntax and morphology*. Oxford, UK: Blackwell and Chicago; University of Chicago Press.

Gábor, B., & Lukács, Á. (2012). Early morphological productivity in Hungarian: Evidence from sentence repetition and elicited production. *Journal of Child Language*, *39*, 411–442. https://doi.org/10.1017/S0305000911000110

Gerken, L. A., Wilson, R., & Lewis, W. (2005). 17-month-olds can use distributional cues to form syntactic categories. *Journal of Child Language*, *32*, 249–268.

Gonzalez-Gomez, N., Schmandt, S., Fazekas, J., Nazzi, T., & Gervain, J. (2019). Infants' sensitivity to nonadjacent vowel dependencies: The case of vowel harmony in Hungarian. *Journal of Experimental Child Psychology*, *178*, 170–183. https://doi.org/10.1016/j.jecp.2018.08.014

Hallé, P. A., Durand, C., & de Boysson-Bardies, B. (2008). Do 11-month-old French infants process articles? *Language and Speech*, *51*, 23–44. https://doi.org/10.1177/00238309080510010301

Höhle, B., & Weissenborn, J. (2003). German-learning infants' ability to detect unstressed closed-class elements in continuous speech. *Developmental Science*, *6*(2), 122–127. https://doi.org/10.1111/1467-7687.00261

Höhle, B., Weissenborn, J., Kiefer, D., Schulz, A., & Schmitz, M. (2004). Functional elements in infants' speech processing: The role of determiners in the syntactic categorization of lexical elements. *Infancy*, *5*(3), 341–353. https://doi.org/10.1207/s15327078in0503_5

Humboldt, W. (1836/1999). *On language. On the diversity of human language construction and its influence on the mental development of the human species*. Cambridge, UK: Cambridge University Press.

Hunter, M. A., & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. *Advances in Infancy Research*, *5*, 69–95.

Ketrez, F. N. (2014). Harmonic cues in speech segmentation: A cross-linguistic corpus study on child-directed speech. *Journal of Child Language*, *41*, 439–461.

MacWhinney, B. (1975). Rules, rote and analogy in morphological formations by Hungarian children. *Journal of Child Language*, *2*, 65–77. https://doi.org/10.1017/S0305000900000891

MacWhinney, B. (1978). *The acquisition of morphophonology*. Monographs of the Society for Research in Child Development 174.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Mahwah, NJ: Erlbaum.

Marquis, A., & Shi, R. (2012). Initial morphological learning in preverbal infants. *Cognition*, *122*, 61–66. https://doi.org/10.1016/j.cognition.2011.07.004

Mikes, M. (1967). Acquisition des catégories grammaticales dans le langage de l'enfant [The acquisition of grammatical categories in child language]. *Enfance*, *20*, 289–298. https://doi.org/10.3406/enfan.1967.2430

Mintz, T. H. (2006). Finding the verbs: Distributional cues to categories available to young learners. In R. M. Golinkoff, & K. Hirsh-Pasek (Eds.), *Action meets word: How children learn verbs* (pp. 31–63). New York, NY: Oxford University Press.

Mintz, T. H. (2013). The segmentation of sub-lexical morphemes in English-learning 15- month-olds. *Frontiers in Psychology*, *4*(24), 1–12. https://doi.org/10.3389/fpsyg.2013.00024

Németh, L. (2003). *A Szószablya fejlesztés [The development of Szószablya]*. Conference on Hungarian Computer Linguistics.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*(1), 27–52. https://doi.org/10.1016/S0010-0277(98)00075-4

Sapir, E. (1921). *Language: An introduction to the study of speech*. New York, NY: Harcourt, Brace and Company.

Schreuder, R., & Baayen, R. H. (1995). Modeling morphological processing. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 131–154). Hillsdale, NJ: Lawrence Erlbaum.

Shafer, V. L., Shucard, D. W., Shucard, J. L., & Gerken, L. A. (1998). An electrophysiological study of infants' sensitivity to the sound patterns of English speech. *Journal of Speech, Language and Hearing Research*, *41*(4), 874–886. https://doi.org/10.1044/jslhr.4104.874

Shi, R., Cutler, A., Werker, J., & Cruickshank, M. (2006). Frequency and form as determinants of functor sensitivity in English-acquiring infants. *The Journal of the Acoustical Society of America*, *119*, EL61-7. https://doi.org/10.1121/1.2198947

Shi, R., & Lepage, M. (2008). The effect of functional morphemes on word segmentation in preverbal infants. *Developmental Science*, *11*(3), 407–413. https://doi.org/10.1111/j.1467-7687.2008.00685.x

Shi, R., Lepage, M., Gauthier, B., & Marquis, A. (2006). Frequency factor in the segmentation of function words in French-learning infants. *The Journal of the Acoustical Society of America*, *119*, 3420. https://doi.org/10.1121/1.4808930

Shi, R., Marquis, A., & Gauthier, B. (2006). Segmentation and representation of function words in preverbal French-learning infants. *Proceedings of the Annual Boston University Conference on Language Development*. 30

Shi, R., & Melançon, A. (2010). Syntactic categorization in French-learning infants. *Infancy*, *15*, 517–533. https://doi.org/10.1111/j.1532-7078.2009.00022.x

Shi, R., Werker, J. F., & Cutler, A. (2006). Recognition and representation of function words in English-learning infants. *Infancy*, *10*(2), 187–198. https://doi.org/10.1207/s15327078in1002_5

Shi, R., Werker, J. F., & Morgan, J. L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, *72*(2), B11–B21. https://doi.org/10.1016/S0010-0277(99)00047-5

Slobin, D. I. (1973). Cognitive prerequisites for the development of grammar. In C. A. Ferguson, & D. I. Slobin (Eds.), *Studies of child language development* (pp. 175–208). New York, NY: Holt, Rinehart & Winston.

Slobin, D. I. (1982). Universal and particular in language acquisition. In E. Wanner, & R. L. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 128–170). Cambridge, UK: Cambridge University Press.

Törkenczy, M. (2011). Hungarian vowel harmony. In M. van Oostendorp, C. J. Ewen, E. Hume, & K. Rice (Eds.), *The Blackwell companion to phonology* (pp. 2963–2990). Malden, MA: Wiley-Blackwell.

How to cite this article: Ladányi E, Kovács ÁM, Gervain J. How 15-month-old infants process morphologically complex forms in an agglutinative language? *Infancy*. 2020;25:190–204. https://doi.org/10.1111/infa.12324